

第三周工作报告

已完成工作

在第三周内，我们小组已经完成的主要工作总结如下：

- 在完整完成了第一次迭代任务的基础上，全面展开了对第二次迭代内容的工作；
- 完成了论文查找所需要的关键词匹配算法，使得在输入关键词进行查找之后，可以依照搜索关键词和数据库中的数据的关联程度(relevance score)来返回排序后的搜索结果；
- 完成了“社区”(Community)板块所需要的后端主体程序，相关API接口说明已经上传到了Github仓库的Wiki中，主要的功能包括发出帖子、进行跟帖、对帖子评论的等功能所需要的数据接口，对应于Panel, Poster, Comment三级内容容器。因为上述内容可能包括非格式化数据，所以全面使用MongoDb作为后台数据库；
- 完成了对于第一次迭代进行的测试时发现的不完善模块的改良；
- 完成了数据可视化进行的元素添加，现在已经包含了包括论文作者、Abstract信息和出版日期等内容，在基于数据库数据作为节点信息行程的力导向图界面中，建立了左中右三个Panel，分别对应于内容的筛选板块、图像展示板块和节点详细信息的说明板块；
- 完成了Community板块中需要的编辑器的制作，分为Markdown和富文本两种模式，允许用户自由选择切换，并支持了编辑内容的暂存功能；
- 对云服务上部署的项目进行了优化，使得在遭遇错误操作时，发生崩溃的机会大大降低；
- 对项目所需要的数据继续进行7 * 24h的爬取，目前数据库中已经有超过24万条论文数据，数据库总的数据容量扩大了一倍。并在同时，对于缺少部分要素的论文和作者数据进行填充；
- 优化了形成力导向图时的算法，考虑到后台数据的庞大规模和对速度的要求，为了兼顾用户体验，默认生成图像时限制最远节点之间的连接度数，在需要扩张图像时再纳入新的节点；
- 针对移动设备的显示效果，对工程前端界面进行了优化。

组内沟通情况

截止目前位置，组内成员功能畅通，在每日会进行小组成员的例行会议，总结成果和布置任务，并对各自负责的任务与队友的任务进行衔接。周五小组进行了有助教全程参与的stand meeting，总结了一周的任务进度，并且布置之后任务。

项目管理情况

针对Github的remote branch和各个组员的local branch，本周的项目管理上出现了一些变化，具体的是：

- 小组新建了**Docs** branch用于对包括进度文档和工作报告在内的文档材料的编写保存；
- 曹金坤同学在**Data** branch上继续进行数据库数据的爬取、清洗、插入等操作；
- 朱锦昊同学的前端工程在阶段性交接之后，从FrontEnd分支转移到了FrontEnd-2分支进行工作；
- 曹金坤同学新建了Editor分支，作为FrontEnd-2分支的下游分支，编写和网站编辑器相关的内容编写，本周已经完成了相关工作，在Merge之后，从FrontEnd-2分支得到了新的下游分支Community，用于对社区板块进行编写；
- 马轲同学继续在BackEnd分支进行后端工程的工作；
- 谢添翼同学继续在PageRank分支进行数据库可视化相关的算法设计和调优工作。

本周遇到的问题

总结上周报告中的问题和回顾本周的工作，有以下问题还需要得到进一步的解决或者关注：

- 包括谢添翼同学的笔记本电脑、曹金坤同学的寝室台式机在内的多台计算机已经被IEEE Xplore网站禁止访问，数据爬取进度因此收到了一些影响；
- 随着数据库容量的增大，包括数据可视化和数据检索在内的操作都遭遇了新的性能瓶颈，需要进一步调优；
- 因为小组租用的云服务器位于海外，而且性能有限，所以在部署工程时，遭遇了一些访问延迟上的限制；

下周工作安排

基于本周的任务进度，对于下周的工作目标，有以下的安排：

- 全面完成对于Community界面的设计和功能实现，包括但不限于：发帖、跟帖、评论等；
- 完成对于数据可视化界面的编写，包括可视化内容的过滤和数据展示等；
- 继续进行数据可用性的优化，使得数据库中paper表中完整数据的数量增加至少3W行；
- 完成Community、数据检索和过滤、数据可视化的过滤在内的网站主体功能对应的后端程序的实现；
- 对上述功能和界面的实现后进行分支的merge，并在网站上进行完整部署；
- 对应于第二次迭代的其他待实现模块。